第13卷第3期 1993年8月
数学研究与评论
JOURNAL OF MATHEMATICAL RESEARCH AND EXPOSITION
Vol. 13 No. 3 Aug. 1993

一般 Gauss—Markov 模型下最小二乘估计是最佳估计的充要条件*

林 春 土
(浙江大学,杭州 310027)

考虑一般的 Gauss—Markov 模型如下

$$M=\{y,X\beta,\sigma^2V\},$$

其中 y 是 $n\times 1$ 的观察值向量,它的期望是 $E(y)=X\beta$,协方差阵是 $D(y)=\sigma^2V$,这里 X 是 $n\times p$ 阶已知矩阵,β 是 $p\times 1$ 的未知参数向量,V 是 $n\times n$ 阶已知的非负定距阵,$\sigma^2>0$ 是未知参数. 本文不须对矩阵 X 和 V 的秩作任何假设.

如果我们用符号 $Py=\mathrm{BLUE}(X_\beta)$ 表示线性统计量 Py 是 $X\beta$ 的最佳线性无偏估计,即在模型 M 下,满足 $E(Py)=X\beta$,而且对一切满足 $E(Gy)=X\beta$ 的 Gy,都有 $D(Py)\leqslant D(Gy)$,这里符号 $A\leqslant B$ 表示 $B-A$ 为非负定的(或记为 $B-A\geqslant 0$). 那么,众所周知,当 $V=I$ 单位阵时,即在模型 $M_0=\{y,X\beta,\sigma^2I\}$ 下,有 $P_Xy=\mathrm{BLUE}(X\beta)$,其中 $P_X=X(X'X)^-X'$,而 A^- 是满足 $AA^-A=A$ 的任意广义逆. 称 P_Xy 为最小二乘估计. 然而,当 $V\neq I$ 时,在一般的模型 $M=\{y,X\beta,\sigma^2V\}$ 下,最小二乘估计 P_Xy 并不是最佳估计. Anderson(1948)最早提出 P_Xy 在一般模型 M 下的优良性问题[1]. 此后,这问题一度成为统计学家研究的热门问题之一,大量的研究论文集中于研究最小二乘估计在 M 下是最佳线性无偏估计的条件. 早期这方面的代表作是 Zyskind(1967)的结果[2],他在这篇文章中总结出在 M 下,$P_Xy=\mathrm{BLUE}(X\beta)$ 的八条等价的充要条件. 以后的许多文献从代数或几何的角度,也给出一些优美的充要条件,例如,[3]—[6]都给出许多独特的结果,值得一提的是 Baksalary 和 Puntanen(1990)的文章[7],他们得出,在 $M=\{y,X\beta,\sigma^2V\}$ 下,$P_Xy=\mathrm{BLUE}(X\beta)$ 的充要条件是如下二条等价的条件之一成立:

(a) $D(P_Xy)\leqslant D(y)$. (b) $r(V-P_XVP_X)=r(V)-r(P_XVP_X)$.

其中 $r(A)$ 表示矩阵 A 的秩. 因为 $D(y)=V$ 和 $D(P_Xy)=P_XVP_X$. 于是,令人感到兴趣的是,结果(a)说明,在模型 M 下,$P_{X}y$ 是否是 $X\beta$ 的 BLUE,无须和一切线性无偏估计量作协方差阵的比较,只要和观察值向量 y 作比较(当然,y 也是一个线性无偏估计);而条件(b)则把最小二乘估计 P_Xy 在 M 下是否最佳问题化为 $D(y)$ 和 $D(P_Xy)$ 二个矩阵之秩的可减性问题. 这样二个结果是解决最小二乘估计优良问题很有启发意义的结果.

本文采用线性子空间 $\mu(X:V)$ 的分解,这里 $\mu(A)$ 表示由矩阵的列向量张成的子空间,对于模型 M,总存在如下分解 $\mu(X:V)=\mu(X:VZ)=\mu(X)\oplus\mu(VZ)$,其中 Z 是满足 $X'Z=0$ 且 $r(Z)=n-r(X)$ 的矩阵,而符号 $\oplus$ 表示线性子空间的直接和. 我们在定理 1 中得到在模型 M 下,P_Xy

* 1991年4月15日收到. 1993年4月14日收到修改稿.

$=$BLUE$(X\beta)$的充要条件为 $\mu(X:V)=\mu(X)\dot{+}\mu(VZ)$，其中符号表示空间的正交和．本文的第二个定理中，我们用关于矩阵 X 的广义逆的种种不变性，来表征最小二乘估计的优良性质．

为了以后证明方便起见，先给如下二个引理：

引理 1 在模型 $M=\{y,X\beta,\sigma^2V\}$ 下，则

$$y\in\mu(X:V),\quad \text{以概率为 1 成立.} \tag{1}$$

证明 对任意的 $n\times1$ 向量 a，如果 $a\perp\mu(X:V)$，那么必有 $a'X=0$，且 $a'V=0$，于是就得到 $E(a'y)=a'X\beta=0$，且 $\mathrm{Var}(a'y)=a'Va=0$．因此，$a'y=0$ 以概率 1 成立，即以概率 1 成立 $y\in\mu(X:V)$．

引理 2 有模型 $M=\{y,X\beta,\sigma^2V\}$ 下，则有

(a) $\mu(X:VZ)=\mu(X)\oplus\mu(VZ)$； (2)

(b) $\mu(X:V)=\mu(X:VZ)$． (3)

证明 如果有 $n\times1$ 向量 $a\in\mu(X)\cap\mu(VZ)$，那么存在向量 t_1 和 t_2，使得 $a=Xt_1=VZt_2$，于是有 $t_2Z'Xt_1=t_2Z'VZt_2=0$，只有 $a=VZt_2=0$．所以 $\mu(Z)\cap\mu(VZ)=\{0\}$．(a)证完．

现在证(b)．显然地，$\mu(X:VZ)\subset\mu(X:V)$，故只要证 $r(X:VZ)=r(X:V)$就可以．因为由(a)，可得 $r(X:V)=r(X)+r[(I-XX^-)V]=r(X)+r(Z'V)=r(X)+r(VZ)=r(X:VZ)$．

由(2)和(3)可见，当 $V=I$ 时，在模型 $M_0=\{y,X\beta,\sigma^2I\}$ 下，有

$$y\in\mu(X:Z)=\mu(X)\dot{+}\mu(Z),\qquad (\text{a.s}). \tag{4}$$

而在一般模型 $M=\{y,X\beta,\sigma^2V\}$ 下，我们得出

$$y\in\mu(X:VZ)=\mu(X)\oplus\mu(VZ),\qquad (\text{a.s}). \tag{5}$$

注意到(4)和(5)之差异是观察值向量 y 的空间分解形式，前者是正交分解，而后者只是直接和分解．然而，在 M_0 之下，成立 $P_Xy=$BLUE$(X\beta)$．下面我们给出更一般的结果．

定理 1 在 $M=\{y,X\beta,\sigma^2V\}$ 下，如下三个条件是等价的：

(a) $P_Xy=$BLUE$(X\beta)$． (b) $\mu(X:VZ)=\mu(X)\dot{+}\mu(VZ)$． (c) $\mu(Z:VX)=\mu(Z)\dot{+}\mu(VX)$．

证明 显然，$E(P_Xy)=X\beta$．因此，在 M 下，$P_Xy=$BLUE$(X\beta)$当且仅当

$$\mathrm{Cov}(P_Xy,Z'y)=P_XVZ\sigma^2=0,$$

易知 $P_XVZ=0\Leftrightarrow X'VZ=0\Leftrightarrow\mu(VZ)\subset\mu(Z)\Leftrightarrow$(b)成立．故(a)$\Leftrightarrow$(b)．

同理，由于 $X'VZ=0\Leftrightarrow Z'VX=0\Leftrightarrow\mu(VX)\subset\mu(X)\Leftrightarrow$(c)成立．定理 1 完全证毕．

为了给出定理 2，再给如下引理 3，其证明见[8]．

引理 3 设 A 和 B 是二个同列矩阵，那么有

$$r\begin{pmatrix}A\\B\end{pmatrix}=r(A)+r[B(I-A^-A)]=r(B)+r[A(I-B^-B)].$$

记 $A\{1,4\}=\{A_m^-:AA_m^-A=A,(A_m^-A)'=A_m^-A\}$．得到 $P_Xy=$BLUE$(X\beta)$的一系列等价条件如下：

定理 2 在 $M=\{y,X\beta,\sigma^2V\}$ 下，那么如下十一个条件是等价的：

(a) $P_Xy=$BLUE$(X\beta)$． (b) $\mu(VX)\subset\mu(X)$． (c)存在矩阵 K，使 $VX=XK$．

(d) $\mu(X:VX)=\mu(X)$． (e) $r(X:VX)=r(X)$．

(f) $XX^-VX=VX$，对任一广义逆 X^- 的选择成立．

(g) $X_m^- VX$ 对每一 $X_m^- \in X\{1,4\}$ 的选择不变.

(h) $r(X_m^- VX)$ 对每一个 $X_m^- \in X\{1,4\}$ 的选择不变.

(i) $r(X_m^- VX)=r(VX)$ 对每一个 $X_m^- \in X\{1,4\}$ 的选择成立.

(j) $\mu[(X_m^- VX)']$ 对每一个 $X_m^- \in X\{1,4\}$ 的选择不变.

(k) $\mu[(X_m^- VX)']=\mu(X'V)$ 对每一个 $X_m^- \in X\{1,4\}$ 成立.

证明 首先(a)$\Leftrightarrow X'VZ=0\Leftrightarrow$(b),而且容易看出条件(b)~(f)都是等价的.

因为对 $X_m^- \in X\{1,4\}$ 的任意选择,$X_m^- X$ 是不变的,故有(c)$\Rightarrow$(g).而且显然又有(g)$\Rightarrow$(h)$\Rightarrow$(j)和(i)$\Leftrightarrow$(k).因此,我们只须证明(h)$\Rightarrow$(i)$\Rightarrow$(c)就足够了.

令 $r(X)=r$,并对 X 作奇异值分解如下

$$X = P\begin{pmatrix} D & 0 \\ 0 & 0 \end{pmatrix}Q', \tag{6}$$

其中 $D=\text{diag}(d_1,d_2,\cdots,d_r)$,$d_i>0(i=1,2,\cdots,r)$ 且 $P'P=I_n$ 和 $Q'Q=I_p$. 易知

$$X_m^- = Q\begin{pmatrix} D^{-1} & E \\ 0 & H \end{pmatrix}P', \tag{7}$$

其中 E 和 H 是任意矩阵.于是我们得到

$$\begin{aligned} r(X_m^- VX) &= r\left(\begin{pmatrix} D^{-1} & E \\ 0 & H \end{pmatrix}P'VX\right) = r\left(\begin{pmatrix} D & 0 \\ 0 & I \end{pmatrix}\begin{pmatrix} D^{-1} & E \\ 0 & H \end{pmatrix}P'VX\right) \\ &= r\left(\begin{pmatrix} I_r & DE \\ 0 & H \end{pmatrix}\begin{pmatrix} A_1 \\ A_2 \end{pmatrix}\right) = r\begin{pmatrix} A_1 + DEA_2 \\ HA_2 \end{pmatrix}, \end{aligned} \tag{8}$$

其中

$$P'VX = \begin{pmatrix} A_1 \\ A_2 \end{pmatrix}. \tag{9}$$

如果(h)成立,令 $E=0,H=0$ 或 $E=0$,由(8)和引理 3,得

$$r(X_m^- VX) = r(A_1) = r\begin{pmatrix} A_1 \\ HA_2 \end{pmatrix} = r(A_1) + r[HA_2(I - A_1^- A_1)]. \tag{10}$$

只有 $HA_2(I-A_1^-A_1)=0$. 因此,$A_2(I-A_1^-A_1)=0$. 再利用引理 3,得到

$$r(A_1) = r(A_1) + r[A_2(I - A_1^- A_1)] = r\begin{pmatrix} A_1 \\ A_2 \end{pmatrix} = r(P'VX) = r(VX). \tag{11}$$

由(10)和(11)式得到(h)$\Rightarrow$(i).

如果(i)成立,令 $E=-D^{-1}A_1A_2^-$ 和 $H=0$. 由(8)式,得到

$$\begin{aligned} r(X_m^- VX) &= r(A_1 - A_1A_2^- A_2) = r[(A_1(I - A_2^- A_2)] \\ &= r\begin{pmatrix} A_1 \\ A_2 \end{pmatrix} - r(A_2) = r(X_m^- VX) - r(A_2). \end{aligned}$$

因此,$A_2=0$. 再由(9)式,得到 $VX = P\begin{pmatrix} A_1 \\ A_2 \end{pmatrix} = P\begin{pmatrix} D & 0 \\ 0 & 0 \end{pmatrix}Q'Q\begin{pmatrix} D^{-1}A_1 \\ 0 \end{pmatrix} = XK$, 其中 $K=Q\begin{pmatrix} D^{-1}A_1 \\ 0 \end{pmatrix}$, (i)$\Rightarrow$(c)证完.

最后,(h)$\Leftrightarrow$(j)和(i)$\Leftrightarrow$(k)都是明显的.因此,定理 2 完全证毕.

注解一 条件(a)和(b)的等价性在早期文献[2]就已经得到.文献[8]中讨论线性模型的

可估性时，首次把可估性和矩阵的广义逆选择的不变性等价起来；文献[7]也对最小二乘的统一理论(Unified Theory)给出矩阵广义逆选择的许多等价的不变性条件．本文的(f)～(k)各条件的等价性，则是首次把 $P_Xy=\mathrm{BLUE}(X\beta)$ 和 X_m^- 的取法不变性相联系起来．

注解二 由条件(b)和(c)．或许容易看出(f)～(k)诸条件的必要性．然而，(f)～(k)各个条件的充分性则是本问题的有趣结果．

注解三 条件(g)～(k)，直观上看，似乎是条件的强弱各异，例如，(g)⇒(h)是显然的，但一般地说，秩的不变性未必有矩阵的不变性．然而，其实质(g)～(k)各条件的不变性却都是等价的．

参 考 文 献

[1] T. W. Anderson(1948), *On the theory of testing serial correlation*, Skand. Aktuarietidskr. 31:88—116.

[2] G. Zyskind(1967), *On canonical forms, non—negative covariance and best and simple least squares linear estimators in linear models*, Ann. Math. Statist. 38:1092—1109.

[3] G. P. H. Styan(1973), *When does least squares give the best linear unbiased estimate? Multivariate Statistical Inference*, North—Holland, Amsterdam, pp. 241—246.

[4] J. K. Baksalary, and R. Kala(1977), *Am extension of a rank criterion for the least squares estimator to be the best linear unbiased estimator*, J. Statist. Plann. Inference, 1:309—312.

[5] S. Puntanen(1987), *On the relative goodness of ordinary least squares estimation in the general linear model*, Acta Univ. Tamper. Ser. A, Vol. 216.

[6] S. Puntanen, and G. P. H. Styan(1989), *The equality of the ordinary least squares estimator and the best linear unbiased estimator(with discussion)*, Amer. Statist., 43, 153—164.

[7] J. K. Baksalary, and S. Puntanen(1990), *Characterizations of the best linear unbiased estimator in the general Gauss—Markov model with the use of matrix partial orderings*, Linear Algebra and Its Applications, 127:363—370.

[8] I. S. Alalouf, and G. P. H. Styan(1979), *Characterizations of estimability in the general linear model*, Ann. of Statist., Vol. 7, No. 1, 194—200.

Necessary and Sufficient Conditions for the Least Squares Estimator to Be the Best Linear Unbiased Estimator in the General Gauss- Markov Model

Lin Chuntu

(Dept. of Math., Zhejiang Univ., Hangzhou)

Abstract

In this paper, we discuss the decomposition of the space $\mu(X:V)$ and the invariance with respect to the choice of a generalized inverse of matrix X in the general Gauss-Markov model. In Theorem 1, we give necessary and sufficient conditions for the least squares estimator $P_Xy=\mathrm{BLUE}(X\beta)$ under the general Gauss-Markov model $M=\{y,X\beta,\sigma^2V\}$. In Theorem 2, we prove that $P_Xy=\mathrm{BLUE}(X\beta)$ under model M and invariant with respect to the choice of a generalized inverse of matrix X are equivalent.